

Transcriptome analysis of apical meristem enriched bud samples for size dependent flowering commitment in *Crocus sativus* reveal role of sugar and auxin signalling

Anjali Chaudhary

CSIR-Institute of Himalayan Bioresource Technology

Kunal Singh

kunal@ihbt.res.in

CSIR-Institute of Himalayan Bioresource Technology

Research Article

Keywords: *Crocus sativus*, Dormancy break, SAUR, Non coding RNA, Flowering primordia

Posted Date: November 28th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3640303/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Molecular Biology Reports on May 3rd, 2024. See the published version at <https://doi.org/10.1007/s11033-024-09574-7>.

Abstract

Background

Cultivation of *Crocus sativus* (saffron) face challenges due to inconsistent flowering patterns and variations in yield. Flowering take place in a graded way with smaller corms being unable to produce flowers. Enhancing the productivity requires a comprehensive understanding of the underlying genetic mechanisms that govern this size based flowering initiation and commitment. Therefore, samples enriched with non-flowering and flowering apical buds from small ($> 6\text{g}$) and large ($< 15\text{g}$) corms were sequenced.

Methods and Results

Apical bud enriched samples from small and large corms were collected immediately after break of dormancy in month of July and performed RNA-sequencing on Illumina platform. *De-novo* transcriptome assembly and analysis using flowering committed buds from large corms at post-dormancy and their comparison with vegetative shoot primordia from small corms pointed out major role of Auxin and ABA hormonal regulation. Many genes with known dual response in flowering development and circadian rhythm like Flowering locus T and Cryptochrome 1 along with a transcript showing homology with small auxin upregulated RNA (SAUR) exhibited induced expression in flowering buds. Thorough prediction of *Crocus sativus* non-coding RNA repertoire has been carried out for the first time. Enolase was found to be acting as a major hub with protein-protein interaction analysis using Arabidopsis counterparts.

Conclusion

Transcripts belong to key pathways including phenylpropanoid biosynthesis, hormone signaling including and carbon metabolism were found significantly modulated. KEGG assessment and protein-protein interaction analysis conform the expression data. Findings unravel the genetic determinants driving the size-based flowering in *Crocus sativus*.

Introduction

Crocus sativus (saffron) is a perennial geophyte herb belonging to the Iridaceae family of monocots. Plant is propagated vegetatively using corms and grown for valuable stigma due to presence of bioactive secondary metabolites with crocin imparting the characteristic color, picrocrocin contributing to the bitter taste, and safranal adding the distinctive fragrance [1]. Saffron is a sterile triploid with a karyotype of $2n = 3x = 24$ and a genome size of $1C = 3.45\text{ Gbp}$ [2, 3]. Saffron is cultivated primarily in Spain, Afghanistan, India, Morocco and Turkey with Iran being the largest producer, accounting for over 90% of the global production [3]. Flowering in *C. sativus* plants is influenced by corm size and weight [4]. The study reported that corms weighing 6–10 g produced an average of 0.45 flowers per corm, while corms

> 38 g weight class yielded approximately 4.5 flowers per corm. Corms less than 6 g size were unable to initiate flowering, resulting in the development of leafy apical primordia and vegetative buds instead [5]. This discrepancy in flowering ability between small and large corms could be attributed to factors of unknown threshold limit of stored nutritional capacity and hormonal regulation. Key factors governing flowering in *C. sativus* and differences in signaling between flowering and non-flowering phenotype are unknown. Transcriptome analysis and expression profiling provided insights into the genes associated with apocarotenoid biosynthesis and production of bioactive compound [6], [7]. Transcript profiling were explored [8] mixing apical buds with axillary buds from different sizes of corm. RNA-Seq analysis of samples immediately before flowering and after flowering identified specific cytochrome P450 genes involved in crocin biosynthesis, the primary pigment in saffron [9]. To assess the molecular mechanism of dormancy break and flower primordia initiation in saffron, very few reports have come out. Plant tissues from large corm as flowering and small corm as non-flowering were compared by transcript expression analysis based on PacBio Iso-seq and Illumina platforms [5] with reported identification of two novel genes PB.20221.2 and PB.38952.1 having significant high expression during flowering. Transcriptomic study of apical bud samples from same size corms for three stages i.e. an undifferentiated period, early flower bud differentiation, and late flower bud differentiation reveal role of many hormone modulated transcripts and sugars [10]. Without specifying the size, different stages of bud transitions was examined with transcriptome analysis [11] reporting induced expression of *FT* (*Flowering locus T*) and *FTIP1* (*FT-Interacting Protein 1*). Yet, none of the work till date have provided comparative transcript expression analysis between flowering apical bud from large corm and vegetatively committed apical bud from small corm immediate after dormancy break. To fill the gap, here we are reporting the transcriptome sequencing and analysis of very early committed stage of flowering bud from large size corms (> 15g) and their comparison with bud from small corms (< 6g, assured being vegetative growth). Analysis revealed role of many genes involved in perception of light and hormones along with SAUR in flowering.

Materials and methods

Plant Material

Corms procured from Kishtwar region, India, were grown in experimental farm field, CSIR-IHBT (latitude 32.0934°N; longitude 76.5439°E) under natural environmental conditions for entire season. Corms were sowed 6 cm deep into furrows with 50 cm distance between each furrow and 6 cm gap between plants. Daughter corms harvested mid-may at season end. Healthy corms were sorted based on size with pool of corms < 6g and > 15g being utilized in the study. Minimum sample pool size for each experiment was 200 (n = 200). Apical bud at July-end stage (S2, L2) was examined for histology using 10x optical microscope (Nikon-BS-S3 olumpus) for each corm sizes. For sampling, the buds were excised using sterile scalpel blade and frozen immediately at -80°C. Samples were collected of two stages of apical bud transitions each, for small and large corm for sequencing at mid-June and July-end.

RNA isolation, Illumina sequencing and assembly

The samples were subjected to RNA extraction using iRIS protocol [12]. Quantity and quality of the extracted RNA was assessed using Bio-analyzer (Agilent Technologies USA) for RIN value, NanoDrop 800 UV-Vis spectrophotometer (Thermo Fisher Scientific, Inc. USA) and TapeStation using HS RNA ScreenTape kit (Agilent). Quality RNA was used for libraries preparation using the NEBNext UltraII RNA library preparation kit, following the manufacturer's instructions. Subsequently, sequencing was performed on the S4 flow cell of NOVASEQ 6000 using 2x159 bp, paired-end chemistry. The raw sequencing data obtained as image files were converted into sequence reads as FASTQ format by CASAVA v.1.8.2. Quality of raw sequencing data was assessed by FastQC (v.0.11.5) and Illumina adapters were removed using Trimmomatic v.0.39 (phred score- 33, quality score- 20, length-20). Filtered reads were assembled-denovo in superTranscripts with Trinity v.2.4.0 using all sequenced samples. All pre-assembled superTranscripts were mapped on available saffron transcriptome data [5] using TransPS. The resultant transcripts were subjected to further clustering using CD-HIT (v.4.6) with a minimum length cutoff of 200 bases. Reads evenness post assembly was assessed by mapping using the BWA-MEM algorithm. BUSCO v.3.1.0, eukaryote_odb10 (Benchmarking Universal Single-Copy orthologs [13]) was employed to evaluate the completeness of the assembly.

Annotation

Transcripts were mapped on blastx-nr mode (v.2.2.29+) with an e-value threshold of e-05. Subsequent annotation was performed on Uniprot (<https://www.uniprot.org/id-mapping>), KEGG (Kyoto Encyclopedia of Genes and Genomes) and planttfdb v.5.0 (Plant Transcription Factor Database) [6]. To classify and predict protein families and domains, the transcripts were scanned against the Interproscan member database signatures.

Identification of protein coding transcripts and expression analysis of S2vsL2 transcriptome data

To differentiate the protein-coding and non-coding expressed RNA sets, Transdecoder was used with minimum ORF length cutoff of 100 amino acids. All such ORFs obtained were scanned against Swissprot and Pfam databases using blastp (version 2.2.29+) and hmmsearch (version 3.3) for positive selection. Separately, total transcripts were also run on CPC2 v.0.1 (Coding Potential Calculator) to predict the coding sequences [14]. Counts of protein-coding expressed transcripts was identified by BWA-MEM algorithm v.0.7.17 (<https://github.com/lh3/bwa>). To assess the profiles of transcripts expressed immediately after dormancy break, transcriptomic data for July end samples i.e. S2, representing vegetative committed bud from small corms and L2, for flowering committed bud from large corms were analysed rigorously using various tools and statistical analysis. DeSeq in R was employed to generate different scatter plots to statistically assessed the transcript expression pattern, their correlation and dispersion by MA plot and dispersion plot. Volcano plot was prepared with log2FC expression data for undeferential, upregulated and downregulated transcripts. ggplot2 was used to prepare the violin plot. DeSeq based values were used for heatmap preparation using RcolorBrewer and gplots with heatmap.2 function to visualize the significant differentially expressed transcripts (DET) with p-value < 0.05 and log2FC > 2.

Gene Ontology terms enrichment, pathway analysis and protein-protein interaction prediction

Gene ontology (GO) enrichment was performed using goatools [15] against the whole transcripts as a background with modulated transcripts as test subject. The multi group bubble plot was prepared using SRPLOT [16]. KASS was used to conduct pathway analysis using KEGG on differentially modulated transcripts [17]. In total 12 plant species were taken for mapping including eight dicots and four monocots viz. *ath* (*Arabidopsis thaliana*), *aly* (*Arabidopsis lyrata*), *bn* (*Brassica napus*), *gmx* (*Glycine max*), *gsj* (*Glycine soja*), *csv* (*Cucumis sativus*), *vvi* (*Vitis vinifera*), *sly* (*Solanum lycopersicum*), *osa* (*Oryza sativa*), *dosa* (*Oryza sativa japonica*), *zma* (*Zea mays*), and *aof* (*Asparagus officinalis*). Default settings were used along with the bi-directional blast. Using the KEGG API (<https://www.kegg.jp/kegg/rest/keggapi.html>), a custom Python script was employed with the request library to fetch Arabidopsis IDs corresponding to specific KO terms. This facilitated data retrieval from the KEGG database. PPI information was obtained from the string database (v.11.5: <https://string-db.org/>) and visual representations of the networks were generated using Cytoscape v.3.9.1 [18]. These visualizations enhanced our understanding of protein relationships within the major pathways.

Non-protein coding transcripts identification and differential expression analysis

All the transcripts that were excluded from Transdecoder and CPC2 were used to identify potential non-coding RNAs including lncRNAs (long non-coding RNAs). C-mi were run to identify probable miRNA precursors. RNAcentral database was searched for all non-coding RNA identification from the assembly. Conserved lncRNAs were predicted by mapping the sequences against the LncPheDB and GreenNC (Green Non coding database) using blastn with an e-value cutoff of $1e-10$.

Quantitative Real Time- PCR (QRT-PCR) for expression validation

RNA was efficiently reverse-transcribed using the High Capacity cDNA synthesis kit (Applied Biosystems, Foster City, CA, USA). The synthesized cDNA was subsequently diluted to a tenfold concentration with DEPC-treated water. All the primers were design using the Primer Express v.3.0.1 (Thermo Fisher Scientific, Waltham, MA, USA), adhering to criteria that included an amplicon size of 80–190 bp, primer lengths of 18–25 nucleotides, a melting temperature (T_m) within the 50–60°C range. Transcript amplification was carried out using Power Sybr Green PCR master mix (ABI) on CFX Opus 96 Real-Time PCR System (BIO-RAD). The chosen QRT-PCR regimen involved an initial denaturation phase at 95°C for 10 minutes, followed by 40 amplification cycles with each cycle consisted of denaturation at 95°C for 15s, annealing at 60°C for 30s, and extension at 72°C for 30s. Tubulin (TUB) and 18s were used as reference genes for normalization. The relative expression data was deciphered using the $2^{-\Delta\Delta CT}$ methodology.

Results

Histological visualization of apical meristem

The saffron plant exhibits a size based gradation of corms with no clearly marked boundary between flowering and non-flowering. Most often, small corms < 6g do not produce flowers and large corms > 12g [19] consistently produce flowers. Figure 1 illustrates the histological variance between vegetatively committed small corms and flowering committed large corm immediate after dormancy break in July-end. Figure 1a-c show undifferentiated leaf primordia in small corms and presence of floscules at the center in large corms (Fig. 1d-f). At least three randomly selected apical buds were microscopically tested and confirmed for flowering primordia in large corms (Fig. 1f) to have higher confidence in downstream transcriptome study. The same was confirmed for vegetatively committed small corms.

Statistics of sequencing data and de-novo assembly

In total ~ 295 million reads obtained from samples consisting of one sample each of small and large corms from mid-June (dormant stage) and second set of samples from July-end (post dormancy break). Total clean reads post Trimmomatic application were 219.57 million with GC content 50–60%. De-novo assembly was performed using all samples for high quality transcript stitching. Assembly provided 2571430 sequence with N-50 of 338 bp. To improve the assembly, superTranscripts were stitched using Trinity by collapsing unique and common sequence regions from splicing isoforms into single linear sequences resulting in 1510262 sequences with N50 of 473. All the sequences obtained till here were mapped to available *C. sativus* dataset by [5], that itself was result of hybrid assembly of PacBio-Illumina, leading to identification of 299581 sequences with N50 of 763. S2 and L2 reads were mapped to the assemble transcripts. Almost, 83.80-88.27% mapping rate was obtained for S2 (small corms: vegetative bud) samples and 82.62–83.09% mapping rate for L2 (Large corm: floral bud). Processing of this transcript pool by removal of contigs having 90% sequence similarity among themselves (considered redundant) and sequence cutoff of 200 bp, finally provided 298870 sequences with N50 of 765 bp. BUSCO analysis using eukaryote lineage of 255 conserved orthologs show presence of 159 out of 255 for full length transcripts while 31.0% mapped as fragmented (79 out of 255) leaving only 6.6% missing (17/255). Therefore, BUSCO analysis suggests the presence of 93.33% eukaryotic conserved orthologs, suggesting that assembly encompasses in-depth transcriptome pool. Application of MISA script identified 18796 short sequence repeats (Supplementary tables S1-S2). Sequencing and assembly pipeline was summarised in Supplementary Fig. 1. Most of the transcripts obtained from assembly were in range of 201–400 bp (Fig. 2a).

Non-Protein coding transcripts identified

All remaining transcripts that were not identified previously as protein coding sequences were categorised as non-coding transcripts (178707) and considered as candidates for identifying LncRNAs and small RNA precursors. In total, 490 transcripts were identified as miRNA precursors by C-mii. A total of 4099 transcripts were mapped on RNA central while GreeNC database provided 745 lncRNAs with

90% identity. Results of RNA-central database suggested Long Noncoding RNA (LncRNA) as predominant non-coding form with 2727 hits, followed by Ribosomal RNA (rRNA) with 1048 transcripts. Miscellaneous RNA (misc_RNA) was catalogued 231 times, other non-coding RNAs (ncRNAs) 28 and small RNA (sRNA) and small nuclear RNA (snRNA) identified with 16 and 15 transcripts, respectively. Eleven transcripts were annotated as signal recognition particle (SRP) while 18 transcripts identified as transfer RNA (tRNA). Run on another database for LncRNA prediction, LncPheDB, documented a total of 1457 entries. Eventually, 3735 non-coding transcripts were identified as the likely LncRNAs (Fig. 2b). Prediction of total LncRNA acting as miRNA precursors was 22 based on the difference between the total hits and unique entries under C-mii.

Many protein coding transcripts differential expressed in floral buds from large corms

Mapping of draft assembly against NR and Uniprot database resulted in 74766 annotated sequences, while 53254, 56230 and 40463 sequences were annotated at KEGG, GO, and PlantTFDB. Transcripts were also scanned against Interproscan database signatures to classify and predict proteins families and domains with more details. A total of 58290 transcripts were reported as potential protein coding transcripts by transdecoder while 14996 were reported by CPC2. Finally, 120163 annotated transcripts were taken as likely protein coding candidates. Many annotation results presented in Supplementary table S3. Expression assessment of protein coding transcripts for S2 vs. L2 stages identified 6112 differentially expressed transcripts (DET). Among them, 338 were induced while 2744 were downregulated in floral bud (L2) relative to vegetative bud (S2). Expression profiles presented in volcano plot showed differential behaviour of two samples after log transformation (Fig. 2c). Transcription factors were among the most modulated transcripts under top 15-up and 15-down-regulated genes based on their expression values (Fig. 2d). Multiple MYB, bZIP and ZF group of transcription factors found place on heatmap.

Gene ontology

Gene Ontology (GO) terms were analysed and categorized (Fig. 2e) into biological processes (BP), molecular function (MF), and cellular component (CC). In total, 674 GO terms were identified for up condition while 576 identified under downregulation. For MF, floral specific terms related to sugar metabolism such as 6-phosphofructokinase activity (GO:0003872), UDP-glucose 6-dehydrogenase activity (GO:0003979) was obtained. Many vegetative growth specific terms encompassing proline biosynthesis (GO:0006561), anatomical structure morphogenesis (GO:0009653), negative regulation of the ethylene-activated signalling pathway (GO:0010105), and leaf vascular tissue pattern formation (GO:0010305) were denoted under downregulated GO terms only under BP category. The enrichment of GO terms after Bonferroni correction method resulted in 15 MF and 2 BP GO terms in total. Major MF GO terms include Zinc ion binding-GO:0008270, Hydrolase activity (acting on ester bonds)- GO:0016788, GO:0140098 and Catalytic activity-GO:0003824 while the BP category include DNA integration-GO:0015074 and Macromolecule metabolic process-GO:0043170 (Fig. 2f).

Many KEGG orthologs representing pathways under modulation revealed

DETs were subjected for KEGG analysis to identify pathways with transcripts having significant modulation. 493 KEGG orthologs (ko) were obtained for transcripts showing upregulation in flowering bud. For downregulated transcripts, a total of 358 ko terms were recognised. Among all the ko identifiers, 373 ko terms were further specific to upregulation, while 238 for downregulated transcripts (Fig. 3a). A total of 120 ko terms were common. Ko categories with the greatest number of transcripts getting modulated were metabolic pathways, biosynthesis of secondary metabolites/amino acids and carbon metabolism (Fig. 3b). Major ko categories ko00720 (carbon fixation in prokaryotes) and ko00710 (carbon fixation in photosynthetic organism) was obtained in carbon metabolism group.

Starch and sugars are major regulators for apical bud destiny

Apart from carbon fixation, carbon metabolism (ko01200) and starch and sucrose metabolism (ko00500) pathways were getting modulated due to significant differential expression of many genes. The genes involved in starch and sucrose metabolism, such as, Glucose-1-phosphate adenylyltransferase (*Cs162717*) and Endoglucanase (*Cs138237*, *Cs171664*, *Cs208287*) were induced more than three folds (log2FC) while soluble Starch synthase (*Cs225848*), 1,4-alpha-glucan branching enzyme (*Cs20486*) and Trehalose-6-phosphate (*Cs230666* and *Cs288059*) expressed more than two folds (> 2FC) in flowering bud (Fig. 3c). Genes encoding enzymes with prior reports of implication in flowering like Phosphoglucomutase (*Cs224419*), Glycogen phosphorylase (*Cs60713*) and Sucrose phosphate phosphatase (*Cs177336*) were detected exclusively in flowering primordia. Three transcripts were observed for Beta-glucosidase enzyme (BGLU) with transcripts *Cs57291* getting downregulated and *Cs153070* and *Cs85501* getting higher expressed in floral bud. Interestingly, Beta-fructofuranosidase (*Cs159238*), and Alpha-amylase (*Cs44494*) were detected in vegetative bud only, while among the two transcripts of sucrose synthase, one (*Cs106601*) was expressed in vegetative bud while other one (*Cs265907*) in floral bud. A heatmap visualising the differential behaviour of these transcripts is presented in Fig. 3d.

Transcripts with role in lignin biosynthesis and flower development were modulated

Many pathways governing secondary metabolites production were found modulated including carotenoid biosynthesis (ko00906), flavonoid biosynthesis (ko00941) and phenylpropanoid biosynthesis (ko00940) when run on KEGG database. Phenylpropanoid pathway take part in forming lignin as end product. General phenylpropanoid pathway genes including Phenylalanine ammonia-lyase (PAL, *Cs174101*), 4-coumarate-CoA ligase (4CL, *Cs50391*) and Shikimate O-hydroxycinnamoyltransferase (HCT, *Cs155054*) were found downregulated (Fig. 4a, b). Expression of Caffeoylshikimate esterase (CSE, *Cs85246*) and Caffeoyl-CoA O-methyltransferase (CCOAMT, *Cs11709*) was detected in floral bud. Lignin

specific pathway genes including Cinnamoyl-CoA reductase (CCR, *Cs195166*), Cinnamyl-alcohol dehydrogenase (CAD, *Cs64672*) and Peroxidase (*Cs4622*) were induced. Caffeic acid 3-O-methyltransferase (COMT, *Cs62963*) was detected in floral bud, specifically. Though reports on involvement of lignin in floral initiation are vague, they must be playing significant role in saffron floral bud initiation.

Genes involved in flowering initiation and signalling such as Timing of CAB expression 1 (TOC1), Gigantea (GI), Cryptochrome 1 (CRY1) and FT were differentially expressed. Among them, (CRY1, *Cs202969*) and FT (*Cs4895*) were detected exclusively in flowering bud while Phytochrome B (PHYB, *Cs2564*), TOC1 (*Cs288652*) and Short vegetative phase (SVP, *Cs283193*) were found downregulated. Gigantea (GI, *Cs75640*) was also 2-fold downregulated. Other key genes involved in regulation of flowering like Suppressor of Constans (SOC1, *Cs193212*), Constans-like (CO; *Cs228654* and *Cs4865*), GA20Ox2 (*Cs156440*), Flowering Control locus A (FCA, *Cs187919*), 14-3-3 (*Cs293409*) and Ultrapetala (*Cs108754*) were induced in flowering primordia.

Hormone signalling contribute to floral determinacy

Many transcripts found differentially expressed under KEGG ortholog term ko0407 (plant hormone signalling) with role in ABA, cytokinin and auxin signalling (Fig. 5a-d). Two genes involved in cytokinin signalling were differentially expressed (Fig. 5a) including Cytokinin receptor histidine kinase 2/3/4 (CRE1, *Cs270630*) and type-B response regulator (ARR-B, *Cs88684*). Genes involved in ABA biosynthesis and signalling were expressed in both samples i.e. vegetative as well as flowering. The Zeaxanthin epoxidase (ZEP, *Cs109205*) expression was observed only in floral apical bud. Xanthoxin dehydrogenase/ABA2 (*Cs2843*) and 9-cis-epoxycarotenoid dioxygenase (NCED, *Cs142472*) were > 2 and > 3 folds downregulated, respectively (Fig. 5b). The (+)-abscisic acid 8'-hydroxylase (CYP707A, *Cs23224*) involved in for ABA degradation was induced 3 times. A transcript showing match with ABA-induced Protein phosphatase 2C (PP2C, *Cs265086*), was downregulated.

Among genes involved in auxin signalling cascades, Auxin influx carrier (AUX1) show both the up (*Cs216224*) and down-regulation (*Cs216221*). In flowering bud, probably, auxin being transported into the cells based on the upregulation (> 5-fold) of Auxin-responsive protein (AUX/IAA; *Cs129866*). The Transport inhibitor response 1 (TIR1; *Cs172274*) which regulate the auxin mediated degradation of IAA-AUX complex, showed 4-fold induction (Fig. 5c, d). Mixed expression behaviour was obtained for three transcripts of Auxin response factors (ARF). One transcript (*Cs14705*) being 2-fold upregulated, second one (*Cs263171*) being 3-fold downregulated and third transcript detected only in small corm (*Cs209884*). Interestingly, key auxin responsive gene Small Auxin Up-regulated RNA (SAUR, *Cs196494*) was among the highest induced transcripts (> 10-fold) in flowering primordia. CsSAUR also mapped in auxin signalling modulation prediction by KEGG and could be acting as key regulator.

Protein-protein interaction reveal Enolase as major hub

Protein-protein interaction (PPI) analysis highlighted strong connections between the sugar(s) and carbon fixation and metabolic pathways. Both the pathway showed significant interactions with the

proteins of phenylpropanoid and carotenoid biosynthetic pathways, suggesting coordinated regulation. PPI network identified few distinct patterns with Enolase (ENO) acting as a central hub with most interactions (Fig. 6a). Though, ENO show limited direct connections with genes belonging to pathways other than carbon fixation, starch and sucrose metabolism. Proteins for carotenoid and flavonoid biosynthesis, hormone signalling and ABC transporters showed notable interactions. DFR (Dihydroflavonol 4-reductase) transcript from flavonoid pathway was predicted to interact with CAD, OMT1 (O-methyltransferase), 4CL1 and PAL1 of the phenylpropanoid pathway and ZEP from the carotenoid pathway. TIR1 known to involve in hormone signalling interacted with CRY1, PHYB, and CCD8 (Carotenoid cleavage dioxygenase 8) suggesting cross-regulation among hormone signalling and light - circadian signalling. A few proteins of brassinosteroid biosynthesis pathway like DWF4 connected with PHYB and CsSAUR analog AT1G16510, while CPD interacted with CRY1, ABA1, and CCD8.

QRT-PCR validated the expression of certain genes involved in growth regulation

QRT-PCR validated the expression dynamics obtained from transcriptomic data (Fig. 6b). Among the genes tested only *CsAGL62* (AGAMOUS like) gave more than 2-fold induction. Transcript expression of almost all the genes including *CsSVP* (SHORT VEGETATIVE PHASE), *CsAP1A* (APETALA1), *CsSEP3.2* (SEPALLATA), *CsAGL6A.1* and *CsAGL80.2* as checked by QRT-PCR were in alignment with RNA-Seq. Transcripts *CsAP1b* and *CsAGL80.2* show partial fluctuation in expression relative to RNA-Seq data, where in former showing almost double the expression under real time while reverse was found for *CsAGL80.2*. To provide more confidence in experiments, two reference genes *Cs18S* and *CsTubulin* were utilised for calculating the fold change in expression in L2 corms relative to S2. Both of them showing at par result, though, expression under tubulin were on higher side.

Discussion

In *C. sativus*, the corm size impacts its ability to flower. Specifically, smaller corms struggle to produce flower, likely due to major nutritional and hormonal needs. Investigations into the specific mechanisms governing this phenomenon contributes to our overall understanding of the regulation of flowering and vegetative growth in saffron. The study identified several pathways related to nutrition being modulated, including carbon fixation, glycolysis, starch and sucrose metabolism. These pathways involved in the conversion and utilization of carbon sources, ultimately influencing their developmental fate. Upregulation of soluble sugar content was found along with high expression of key genes involved in sugar metabolism in the flowering bud of saffron [10]. Many genes like Trehalose-6-phosphatase/synthase (TPP/TPS), PYG and endoglucanase with role in providing free sugars (glucose, Trehalose and sucrose) were found upregulated in floral apical bud in our work. TPS facilitate production of Trehalose and Trehalose-6-phosphate (T6P) signalling molecules for carbohydrate allocation with significant role in flower induction [20]. Furthermore, T6P influences the expression of FT gene accelerating the floral transition [10]. Significance of sugar metabolism, T6P signalling, and the interplay of various genes in regulating the floral transition in other plants has been emphasised [21]. As many

results were in accordance with reports in other plant species, they highlight the crucial role of sugar signalling in flower induction [22], [23].

Along with carbon and sugar based signalling many transcripts involved in hormone mediated signalling also showed modulation. Downregulation of 9-cis-epoxycarotenoid dioxygenase (NCED) in floral bud a crucial enzyme in abscisic acid (ABA) biosynthesis indicated that ABA might be promoting vegetative bud formation rather than floral bud. Decrease in ABA levels was reported during the flowering induction accompanied by changes in the expression of genes involved in ABA biosynthesis and signal transduction, such as NCED, PP2C and ABF [10]. Negative regulation of flowering in saffron by ABA was further strengthened by upregulation of CYP707A, a gene involved in ABA degradation and downregulation of PP2C. The PP2C themselves was downregulated. Though, ABA is not universal negative regulator of flowering as litchi plant (*Litchi sinensis*) require ABA influence for floral induction [24]. It was suggested that ABA interact with sugars and participates in sugar metabolism through SnRK1. One more report suggest negative regulatory role of ABA for flower induction in saffron [11] as decrease in hormone level was observed under flowering. Therefore, The current study reinforced the concept of negative ABA regulation in saffron.

Similarly, Auxin and cytokinin signalling pathways play a major role in floral fate determination. Few transcripts related to cytokinin signalling were significantly modulated. Like, the expression of cytokine receptor *CRE1* was observed in L2 condition only. The most promising modulated pathway for floral morphotype could be auxin signalling. The dynamics of PIN1-mediated polar auxin transport and the biosynthesis of auxins in the meristem's inner regions have attracted significant attention in the past [11]. Presence of auxin efflux carriers (PIN1) was detected in floral bud. Additionally, expression of a gene encoding an auxin-influx carrier (AUX1) was modulated along with the upregulation of CsSAUR with both of them being early responsive genes involved in auxin-mediated responses. Role of auxin was emphasized in the regulation of the flowering transition in saffron [10]. They observed significant differences in the levels of auxin and several auxin-related genes, including AUX1, ARF, SAUR, and AUX/IAA, between the dormant stage (DS) and the flowering stage (FS). Furthermore, it was reported that the IAA-inducible protein SAUR acts to inhibit the activity of the PP2C-type protein phosphatase, thereby influencing plant growth and development in *Rosa rugosa* [25]. Their study revealed the up-regulated expression of *RrSAUR* and the down-regulated expression of *RrPP2C* from dormant to initially differentiated floral bud, suggesting a potential negative regulation of *RrPP2C* by *RrSAUR* during the flowering transition process. Therefore, intricate regulatory network interplay between IAA and ABA signalling pathways was present in floral development similar to current report.

Other than hormones, genes having role in lignin synthesis and phenylpropanoid pathway such as COMT, CCR, CAD, and CSE were getting upregulated suggesting the importance of lignification during floral initiation in saffron. Research on leaf carotenoid metabolism highlights lutein's vital role [26]. The environment effects on lutein production in leaves was studied and reported its importance in light capture and plant protection. The study showed that lutein was more prevalent in leaves than in flowers. This might explain the specific expression of LUT1 in vegetative buds only. Due to limited genomic

information available on saffron, a PPI network was constructed with help of *Arabidopsis thaliana* orthologs. PPI suggested a few major hub among many interaction including Enolase. Appearance of a SAUR ortholog in interaction, further raises its importance and could be a major factor in saffron for flowering

Our study addresses a notable limitation from previous reports as comparison of a vegetative small corm bud with flowering bud from a large size corm has been carried first time, here. Earlier reports predominantly focused on floral development and differentiation at late stage or comparison of transition with in same size-weight corms. Most often, samples were subjected to cold treatment or derived from auxiliary buds of flowering phenotypes, potentially retaining the cues. Additionally, TIR1, SnrK1, G-6-P were not reported earlier under floral transition [10] and FT, SOC1 [5]. Detection of these and many other transcripts with differentially expression fill the lacunae for such reports.

Conclusion

The process of flowering in *C. sativus* is influenced by corm size (weight). Small corms exhibit a limited capacity towards flowering likely due to difference in sugar and starch signaling that may based on storage capacity of nutrition in large corms. Many long non-coding RNAs were identified from saffron first time. Pathways analysis indicated modulation of transcripts belonging to Auxin, Cytokinin and ABA signaling along with phenylpropanoid, flavonoid and carotenoid pathways. *CsSAUR* identified as one of the key factors for flowering promotion in saffron large corm. Further investigations would facilitate about our understanding of how auxin determine the fate of the meristem in the main bud and their interactions with other phytohormones and signalling pathways.

Declarations

Data Availability

Data will be shared on request.

Funding

Work was supported by CSIR project MLP0168 (MLP0049).

Author contributions

KS and AC designed and conceptualized the work. AC executed the experiments. AC and KS analyzed the data and wrote the manuscript. KS arranged funding.

Acknowledgment

AC is thankful to CSIR, India for fellowship.

Competing Interest

The authors declare no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

All authors given their consent to publish

References

1. Brandizzi F, Grilli Caiola M (1998) Flow cytometric analysis of nuclear DNA in *Crocus sativus* and allies (Iridaceae). *Plant Systematics and Evolution* 211:149 – 54.
<https://doi.org/10.1007/BF00985356>
2. Schmidt T, Heitkam T, Liedtke S, Schubert V, Menzel G (2019) Adding color to a century-old enigma: Multi-color chromosome identification unravels the autotriploid nature of saffron (*Crocus sativus*) as a hybrid of wild *Crocus cartwrightianus* cytotypes. *New Phytol* 222(4):1965–1980.
<https://doi.org/10.1111/nph.15715>
3. Busconi M, Wischnitzki E, Del Corvo M, Colli L, Soffritti G, Stagnati L, Fluch S, Sehr EM, de los Mozos Pascual M, Fernández JA (2021) Epigenetic variability among saffron crocus (*Crocus sativus* L.) accessions characterized by different phenotypes. *Front Plant Sci* 12:642631.
<https://doi.org/10.3389/fpls.2021.642631>
4. Douglas MH, Smallfield BM, Wallace AR, McGimpsey JA (2014) Saffron (*Crocus sativus* L.): The effect of mother corm size on progeny multiplication, flower and stigma production. *Sci Hort* 166:50–58. <https://doi.org/10.1016/j.scienta.2013.12.007>
5. Qian X, Sun Y, Zhou G, Yuan Y, Li J, Huang H, Xu L, Li L (2019) Single-molecule real-time transcript sequencing identified flowering regulatory genes in *Crocus sativus*. *BMC Genomics* 20:1–8.
<https://doi.org/10.1186/s12864-019-6200-5>
6. Baba SA, Mohiuddin T, Basu S, Swarnkar MK, Malik AH, Wani ZA, Abbas N, Singh AK, Ashraf N (2015) Comprehensive transcriptome analysis of *Crocus sativus* for discovery and expression of genes involved in apocarotenoid biosynthesis. *BMC Genomics* 16:1–4. <https://doi.org/10.1186/s12864-015-1894-5>
7. Jain M, Srivastava PL, Verma M, Ghangal R, Garg R (2016) De novo transcriptome assembly and comprehensive expression profiling in *Crocus sativus* to gain insights into apocarotenoid biosynthesis. *Sci Rep* 6(1):22456. <https://doi.org/10.1038/srep22456>
8. Sharma M, Kaul S, Dhar MK (2019) Transcript profiling of carotenoid/apocarotenoid biosynthesis genes during corm development of saffron (*Crocus sativus* L). *Protoplasma* 256:249–260.
<https://doi.org/10.1007/s00709-018-1296-z>

9. Gao G, Wu J, Li B, Jiang Q, Wang P, Li J (2021) Transcriptomic analysis of saffron at different flowering stages using RNA sequencing uncovers *cytochrome P450* genes involved in crocin biosynthesis. *Mol Biol Rep* 48:3451–3461. <https://doi.org/10.1007/s11033-021-06374-1>
10. Hu J, Liu Y, Tang X, Rao H, Ren C, Chen J, Wu Q, Jiang Y, Geng F, Pei J (2020) Transcriptome profiling of the flowering transition in saffron (*Crocus sativus* L). *Sci Rep* 10(1):9680. <https://doi.org/10.1038/s41598-020-66675-6>
11. Renau-Morata B, Nebauer SG, García-Carpintero V, Canizares J, Minguet EG, De los Mozos M, Molina RV (2021) Flower induction and development in saffron: Timing and hormone signalling pathways. *Ind Crops Prod* 164:113370. <https://doi.org/10.1016/j.indcrop.2021.113370>
12. Ghawana S, Paul A, Kumar H, Kumar A, Singh H, Bhardwaj PK, Rani A, Singh RS, Raizada J, Singh K, Kumar S (2011) An RNA isolation system for plant tissues rich in secondary metabolites. *BMC Res Notes* 4(1):1–5. <https://doi.org/10.1186/1756-0500-4-85>
13. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM (2018) BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol* 35(3):543–548. <https://doi.org/10.1093/molbev/msx319>
14. Kang YJ, Yang DC, Kong L, Hou M, Meng YQ, Wei L, Gao G (2017) CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res* 45(W1):W12–W16. <https://doi.org/10.1093/nar/gkx428>
15. Jiao Y, Hu Q, Zhu Y, Zhu L, Ma T, Zeng H, Zang Q, Li X, Lin X (2018) GOATOOLS: A Python library for Gene Ontology analyses. *Sci Rep* 8(1):10872. <https://doi.org/10.1038/s41598-018-28948-z>
16. Tang D, Chen M, Huang X, Zhang G, Zeng L, Zhang G, Wu S, Wang Y (2023) SRplot: A free online platform for data visualization and graphing. *PLoS ONE* 18(11):e0294236. <https://doi.org/10.1371/journal.pone.0294236>
17. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35(suppl2):W182–W185. <https://doi.org/10.1093/nar/gkm321>
18. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–2504. <https://doi.org/10.1101/gr.1239303>
19. Negbi M, Dagan B, Dror A, Basker D (1989) Growth, flowering, vegetative reproduction and dormancy in the saffron crocus (*Crocus sativus* L). *Isr J Bot* 38(2–3):95–113. <https://doi.org/10.1080/0021213X.1989.10677116>
20. Wahl V, Ponnu J, Schlereth A, Arrivault S, Langenecker T, Franke A, Feil R, Lunn JE, Stitt M, Schmid M (2013) Regulation of flowering by trehalose-6-phosphate signaling in *Arabidopsis thaliana*. *Science* 339(6120):704–707. <https://doi.org/10.1126/science.1230406>
21. Wang JW (2014) Regulation of flowering time by the miR156-mediated age pathway. *J Exp Bot* 65(17):4723–4730. <https://doi.org/10.1093/jxb/eru246>

22. Xing LB, Zhang D, Li YM, Shen YW, Zhao CP, Ma JJ, An N, Han MY (2015) Transcription profiles reveal sugar and hormone signaling pathways mediating flower induction in apple (*Malus domestica* Borkh). Plant Cell Physiol 56(10):2052–2068. <https://doi.org/10.1007/s11103-018-0801-2>
23. Ortiz-Marchena MI, Romero JM, Valverde F (2015) Photoperiodic control of sugar release during the floral transition: What is the role of sugars in the florigenic signal? Plant Signal Behav 10(5):e1017168. <https://doi.org/10.1080/15592324.2015.1017168>
24. Cui Z, Zhou B, Zhang Z, Hu Z (2013) Absciscic acid promotes flowering and enhances *LcAP1* expression in *Litchi chinensis* Sonn. South Afr J Bot 88:76–79. <https://doi.org/10.1016/j.sajb.2013.05.008>
25. Wang X, Zhao F, Wu Q, Xing S, Yu Y, Qi S (2023) Physiological and transcriptome analyses to infer regulatory networks in flowering transition of *Rosa rugosa*. Ornam Plant Res 3(1):1–2. <https://doi.org/10.48130/OPR-2023-0004>
26. Dhami N, Cazzonelli CI (2020) Environmental impacts on carotenoid metabolism in leaves. Plant Growth Regul 92(3):455–477. <https://doi.org/10.1007/s10725-020-00661-w>

Figures

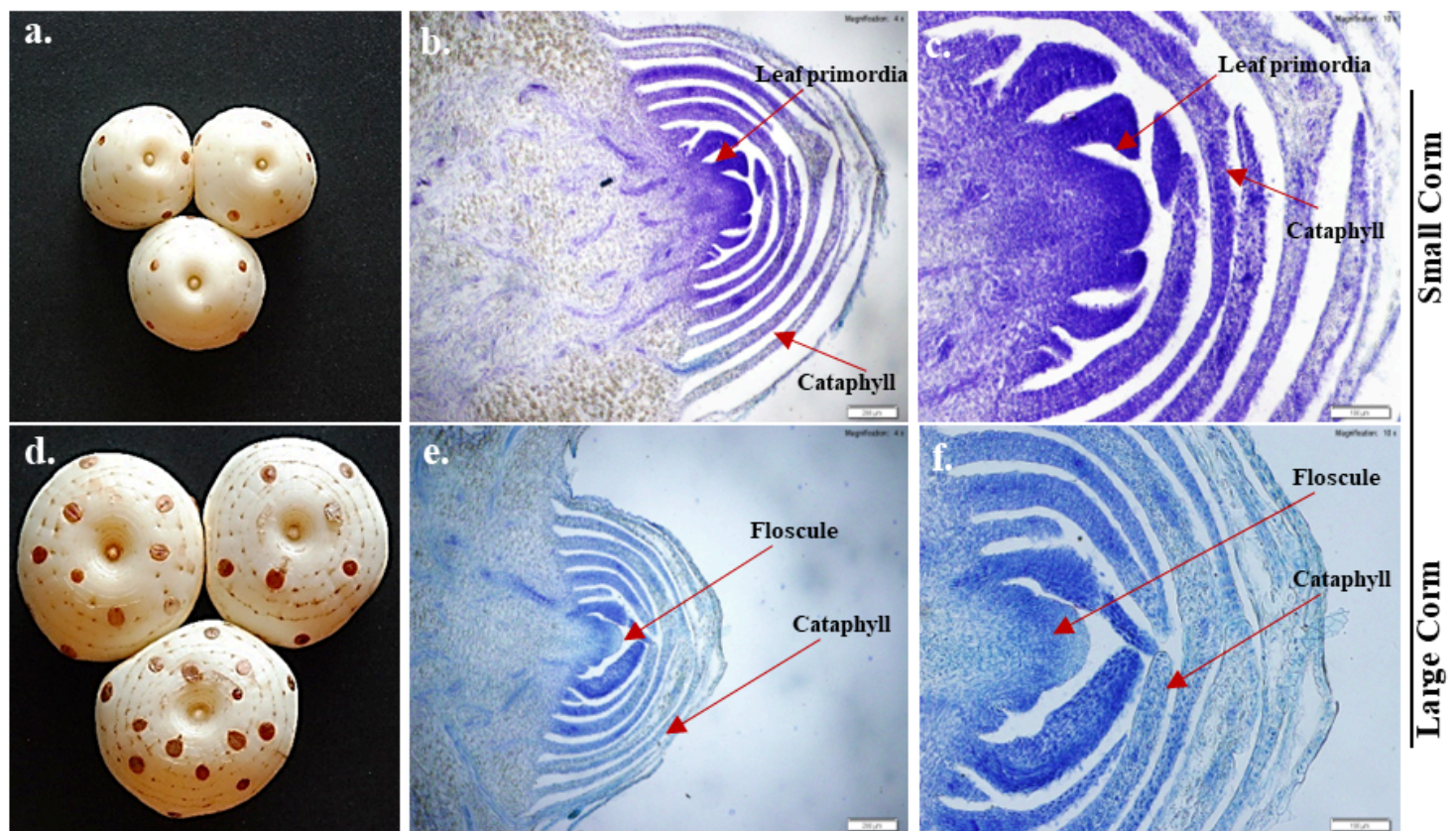


Figure 1

(a) Small corm (< 6 g) without cataphyll and tunics (b) Light microscopy image of small corm apical bud at 4x magnification (c) at 10x magnification (d) Large corm (> 14g) without tunics and cataphyll (e) Light microscopy image of large corm apical bud at 4x magnification (c) at 10x magnification

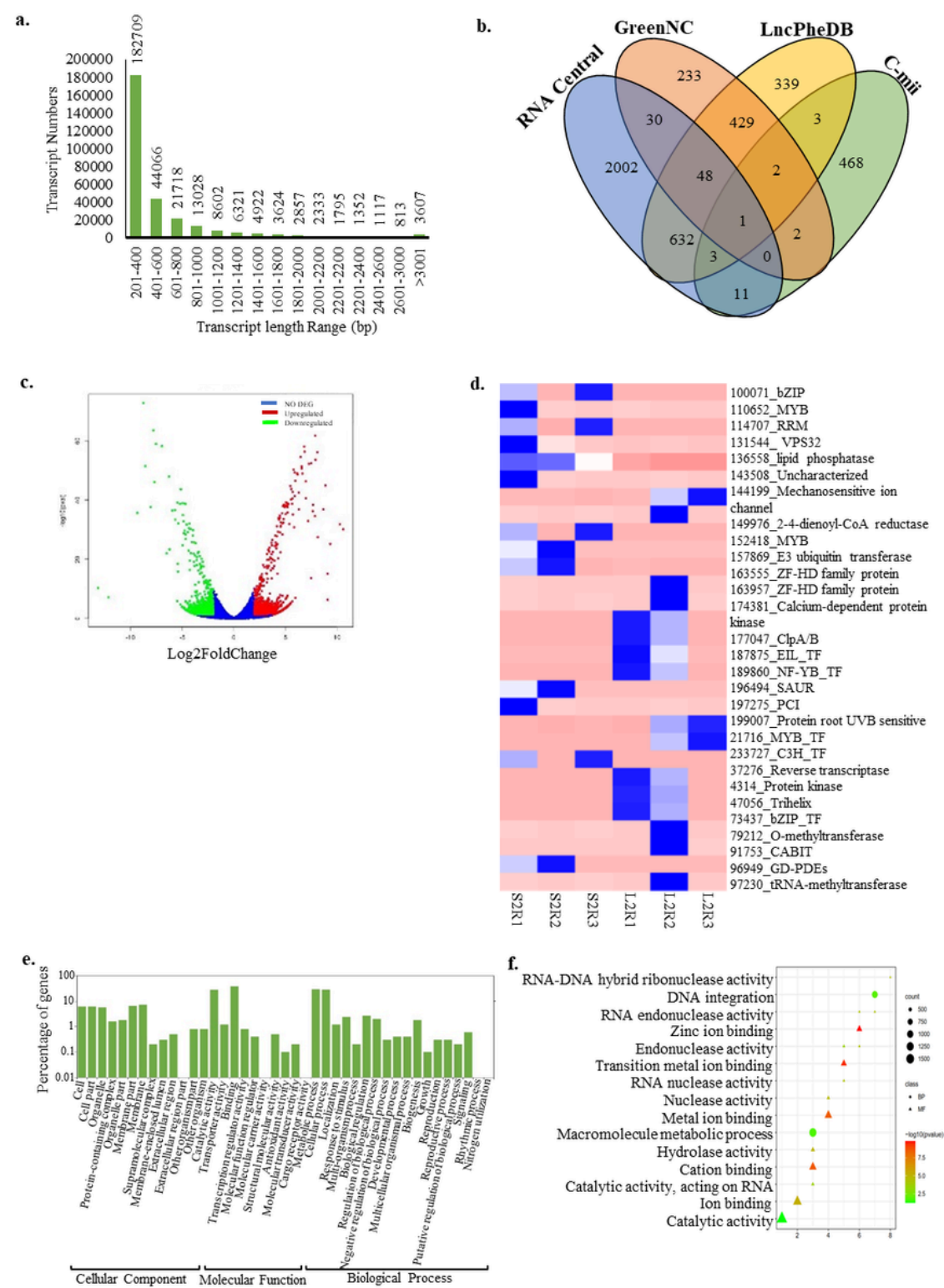
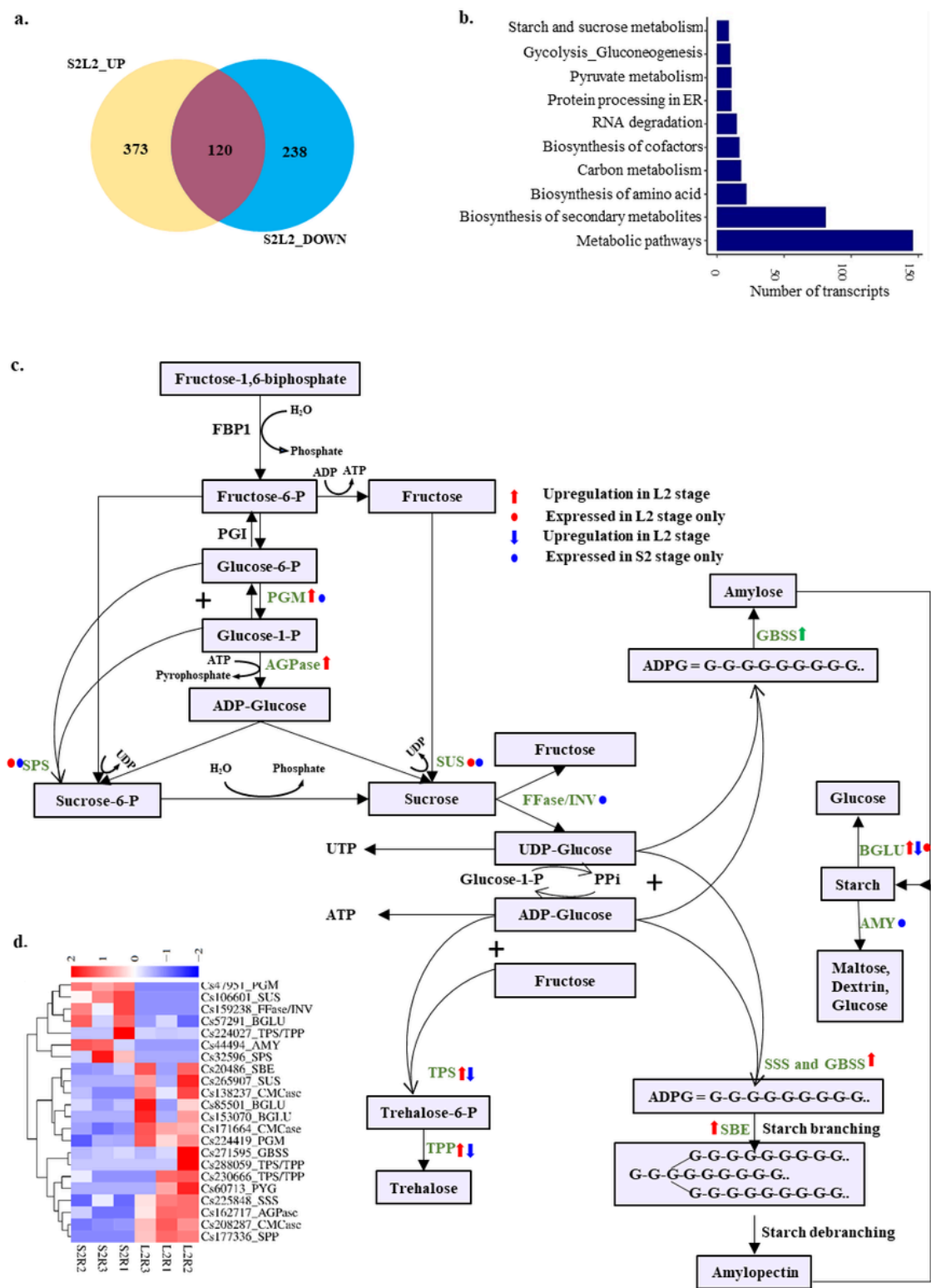


Figure 2

(a) Length distribution of transcripts after assembly (b) Venn diagram representing results of various tools used for non-coding RNA- LncRNA identification. (c) Volcano plots showing overall differential

pattern of transcripts for S2 vs. L2 comparison (d) Heatmap of top 30 differentially expressed transcripts (15-up and 15-down) with presence of *CsSAUR* and many saffron transcription factors. (c) Gene Ontology (GO) terms distribution with (d) GO enrichment plot depicting major categories like catalytic and hydrolase activity and ion binding.



(a) Venn diagram representing number of up and down regulated transcripts based on KEGG orthologs (ko) based significance analysis, and (b) Plot representing the distribution of KEGG enriched major pathways (c) Analysis and representation of starch and sucrose metabolism pathway depicting major transcripts getting modulated (Abbreviation: FBP1- Fructose-1,6-bisphosphatase1; PGI- Phosphoglucose isomerase/Phosphoglucoisomerase; PGM- Phosphoglucomutase; AGPase- ADP-glucose pyrophosphorylase; SPS- Sucrose-phosphate synthase; SUS- Sucrose synthase; FFase: Beta-fructofuranosidase (aka INV- Invertase); SSS- Soluble starch synthase; GBSS- Granule-bound starch synthase; AMY- Alpha-amylase; BGLU- Beta-glucosidase; TPS- Trehalose-6-Phosphate synthase; TPP- Trehalose-6-phosphatase; SBE- Starch branching enzyme; CMCase- Carboxymethylcellulases (Endoglucanase); PYG- Glycogen phosphorylase). Green fonts indicate the modulated transcripts. (d) heatmap representing differential expression of major transcripts from pathway.

esterase; CAD- Cinnamyl alcohol dehydrogenase; COMT- Caffeic acid 3-O-methyltransferase/acetylserotonin -O-methyltransferase; PRX- Peroxidase; CCoAMT- Caffeoyl-CoA O-methyltransferase). Green fonts indicate the modulated transcripts. (b) heatmap representing differential expression of major transcripts from pathway.

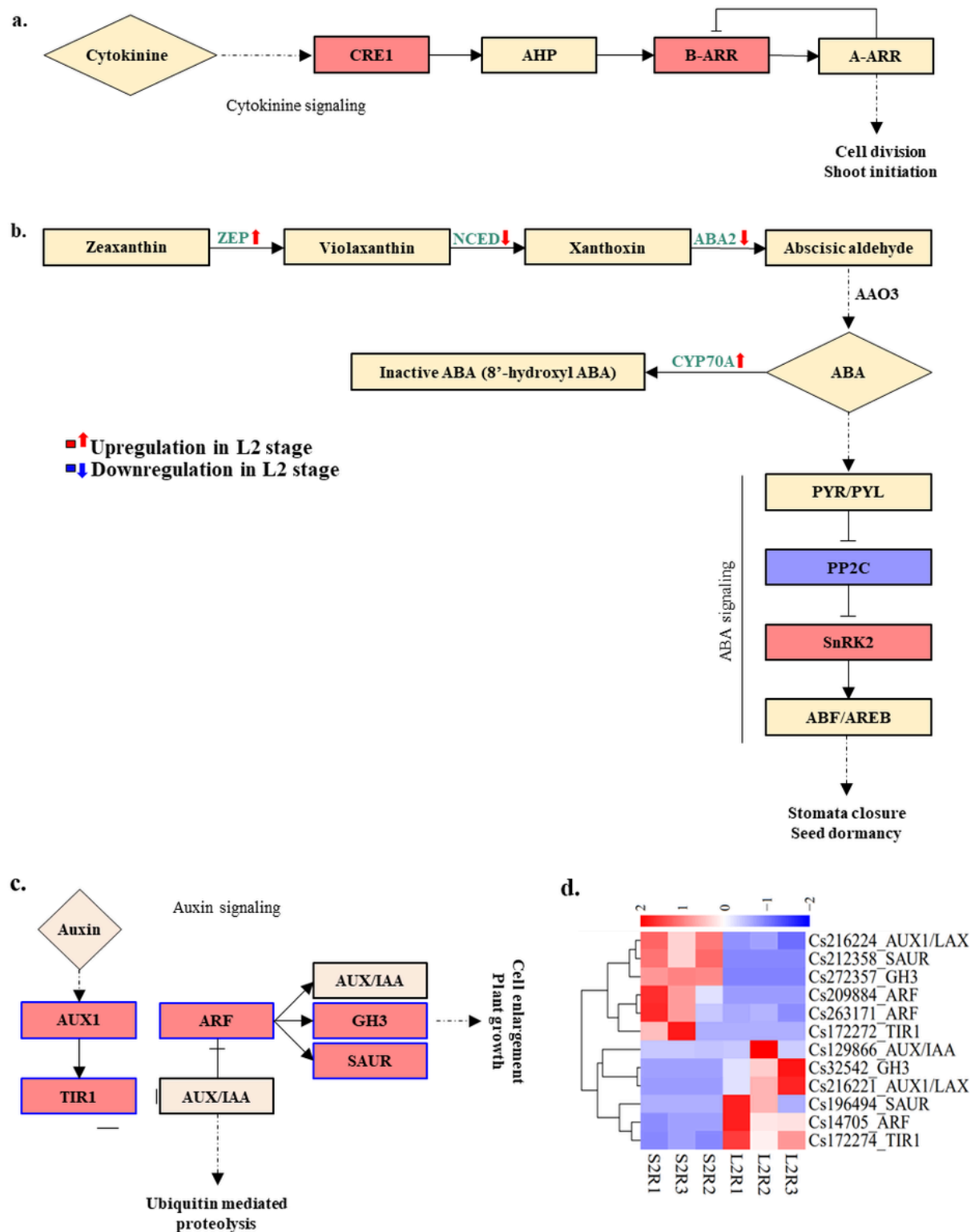
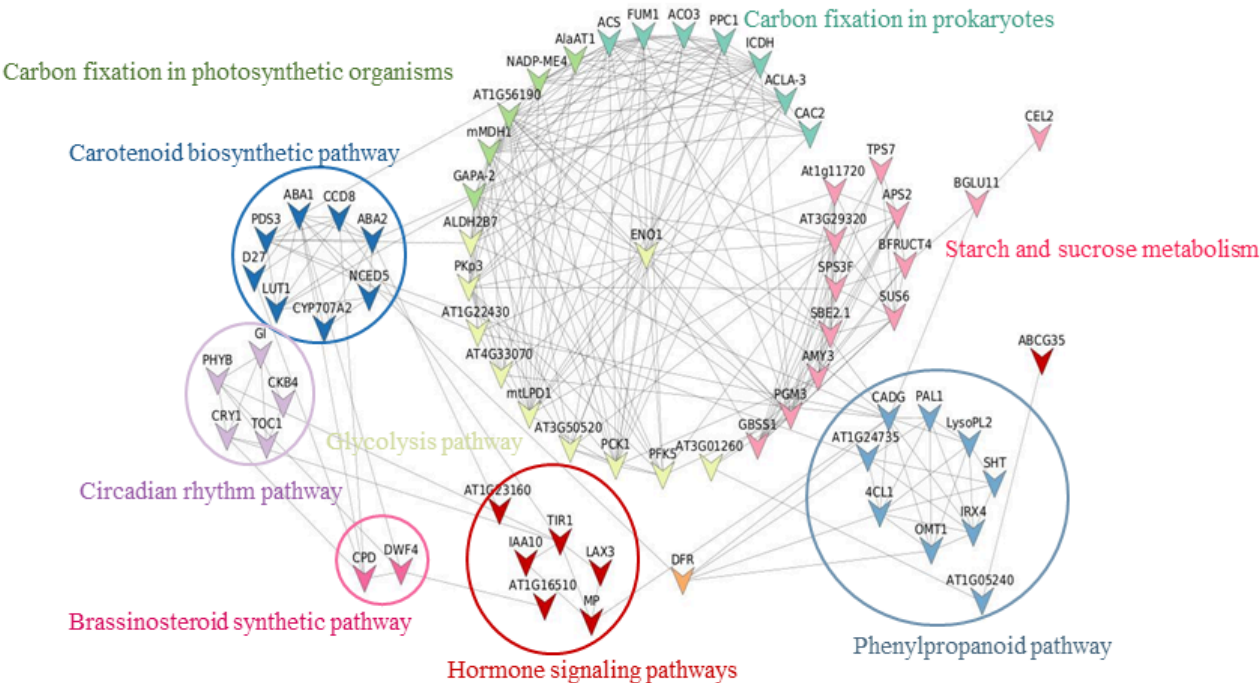


Figure 5

Model of major hormone signalling cascades involved in flowering regulation including (a) Cytokinin signalling (b) ABA biosynthesis and signalling (c) Auxin signalling pathway.

(Abbreviation: CRE1- Cytokinin receptor 1; B-ARR- Two-component response regulator ARR-B family; ZEP- Zeaxanthin epoxidase; NCED- 9-cis-epoxycarotenoid dioxygenase; ABA2- Xanthoxin dehydrogenase; AAO3- Absciscic aldehyde oxidase 3; CYP707A1- (+)-absciscic acid 8'-hydroxylase; PP2C- Protein phosphatase2C; SnRK2- serine/threonine-protein kinase). (d) Visual representation of major differentially expressed transcripts under auxin signalling cascade like *CsTIR1*, *CsSAUR* and *CsAUX* by heatmap.

a.



b.

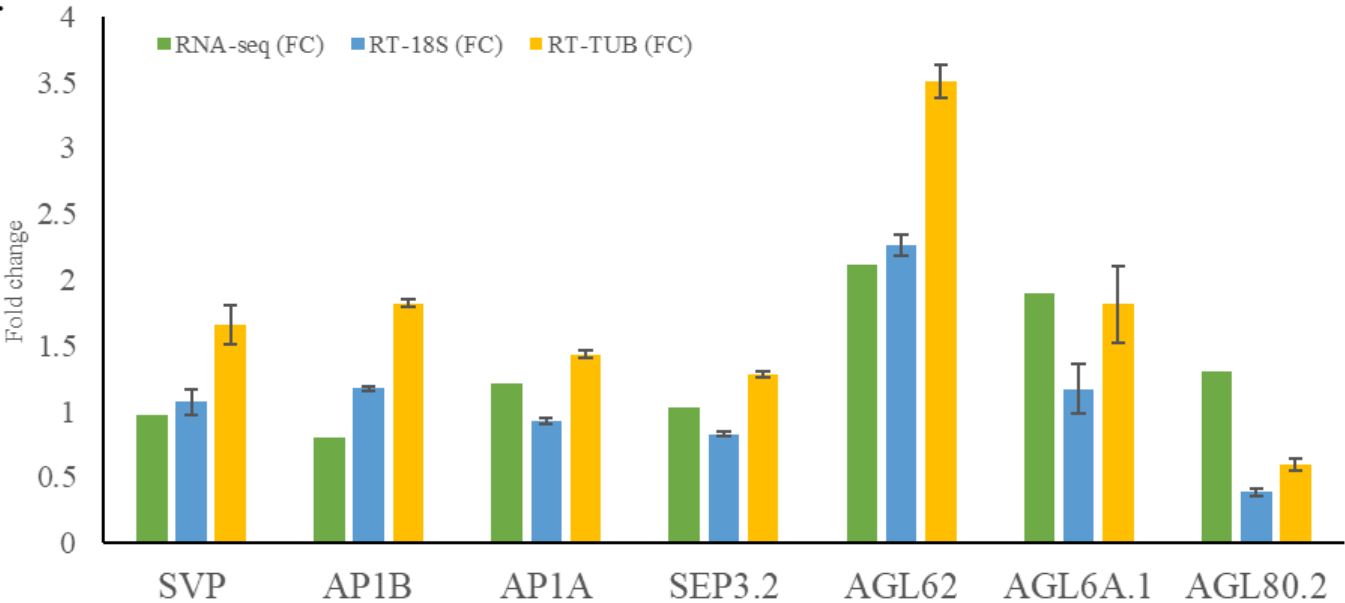


Figure 6

(a) Protein-Protein Interaction network of modulated pathways with major hubs including Enolase. (b) Validation of RNA-sequencing (transcriptome) data with quantitative real time PCR showing similar expression pattern for all the genes tested. Analysis represents expression data under two reference genes for normalization.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supple.fig.s1.pdf](#)
- [SupplimentaryTables.pdf](#)